Exploring the EGF/Ras and Notch/DSL Signaling Pathways in Members of the Caenorhabditis Genus

Research Thesis

Presented in partial fulfillment of the requirements for graduation *with research distinction* in the undergraduate colleges of The Ohio State University

by

Victoria Simons

The Ohio State University
April 2016

Project Advisor: Professor Helen Chamberlin, Department of Molecular Genetics

Introduction

Cell signaling pathways are a key component to the development and growth of nearly all organisms; these paths are so vital that there is often high conservation in them across even vastly different species. While a variety of model systems are used worldwide to study the role of these paths in development and various diseases, this project focuses on nematode species, specifically nematodes of the Caenorhabditis genus. These organisms provide a simple system for the study of how important signaling pathways (such as EGF/Ras) influence cell growth and division. Comparison among these species allows for a better understanding of how different genetic backgrounds influence these processes. This is because, although they are closely related, these worm species exhibit some differences in how they respond to certain stressors, like chemical inhibitors and loss of genes. These differing responses can be likened to different responses among humans to drug treatments: where a drug might cure a certain part of the population, another part may have no reaction to it at all. To better understand both these signaling pathways and these organisms, it is important to understand how they have changed from an evolutionary perspective; over time, genes can duplicate, get deleted, or display new variations in splicing. Changes to the genome of an organism are how they evolve over time, and by examining these processes in more detail, we can learn more about how members of Caenorhabditis are related, and how EGF/Ras and Notch/DSL cope with changes to their key genes.

Signal transduction pathways are vital components of a functioning organism. By a carefully controlled network of ligands and receptors, cells can “talk” to each other and communicate what fate to adopt, and how to react to stimuli. The Ras/MAPK and Notch/DSL signaling transduction pathways play a key role in cancer development in humans; in nematodes,

Figure 1

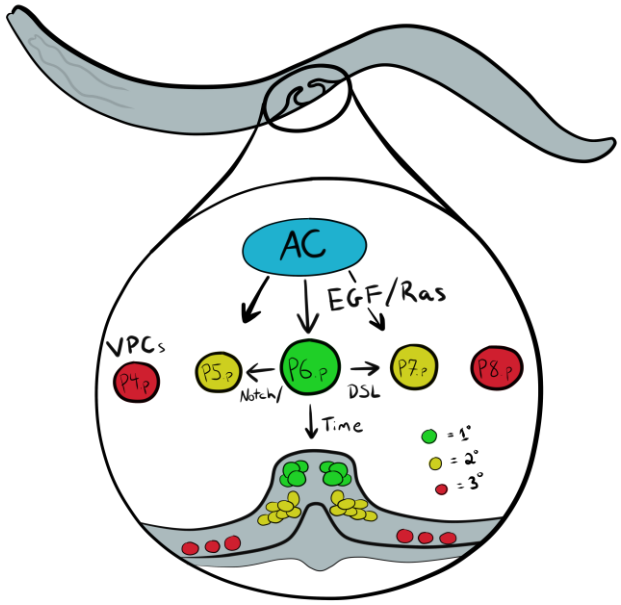
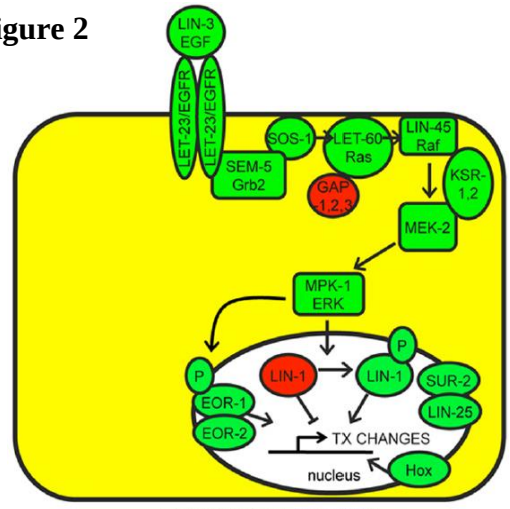


Figure 1. The process of signaling vulvagenesis from the anchor cell, and subsequent cell divisions.

the pathways help with a crucial step of worm growth: development of the vulva tissue from precursor cells. Three of six precursor cells on the ventral surface of the worm divide to make 22 mature vulva cells. A seventh cell – termed the anchor cell (AC) – is dorsal to the vulva precursor cells (VPCs) and sends the necessary signal for the cells to divide. In a normal worm three of the six cells; P.3, P.4, and P.8; adopt a tertiary (3^o) fate, which allows them to fuse with surrounding epidermal tissue. The remaining three undergo additional cell divisions, with P.5 and P.7 taking on a secondary (2^o) fate, creating the walls of the vulva and attaching to the epidermis via the tertiary cells, and P.6, adopting primary fate (1^o), and contributing to the top of the vulva. This concept is demonstrated visually in **Figure 1**.

This project sought to study the EGF/Ras pathway (**Figure 2**) and the Notch/DSL pathway (**Figure 3**), especially the diversity of each pathway, and how they work together to allow successful vulvagenesis. Since both these pathways are crucial to a nematode's developing vulva, it is important that no errors occur in development, otherwise the worm may

Figure 2



LET-23/EGFR-Ras-ERK *(Sundram, 2013)*

Figure 2. The EGF/Ras signal cascade. *Lin-3* is the ligand that starts the process. Figure adapted from Sundram (2013), available on WormBook.org

not be able to properly lay its eggs and generate the next generation of worms. Therefore there was particular interest in how these pathways have developed evolutionarily over time, and whether there was a difference in gene duplication and gene splicing between the two paths, and among different Caenorhabditis members. Experimental work focused on the signaling protein LIN-3/EGF in the EGF/Ras pathway, specifically examining differences between the gene in *C. briggsae* and *C. elegans* worms, but also examining splicing in *C. brenneri*, *C. remanei*, and *C. japonica*. The intent of this experimental work is to broaden the knowledge and understanding currently available for *lin-3* in Caenorhabditis, particularly in terms of gene splicing.

Figure 3

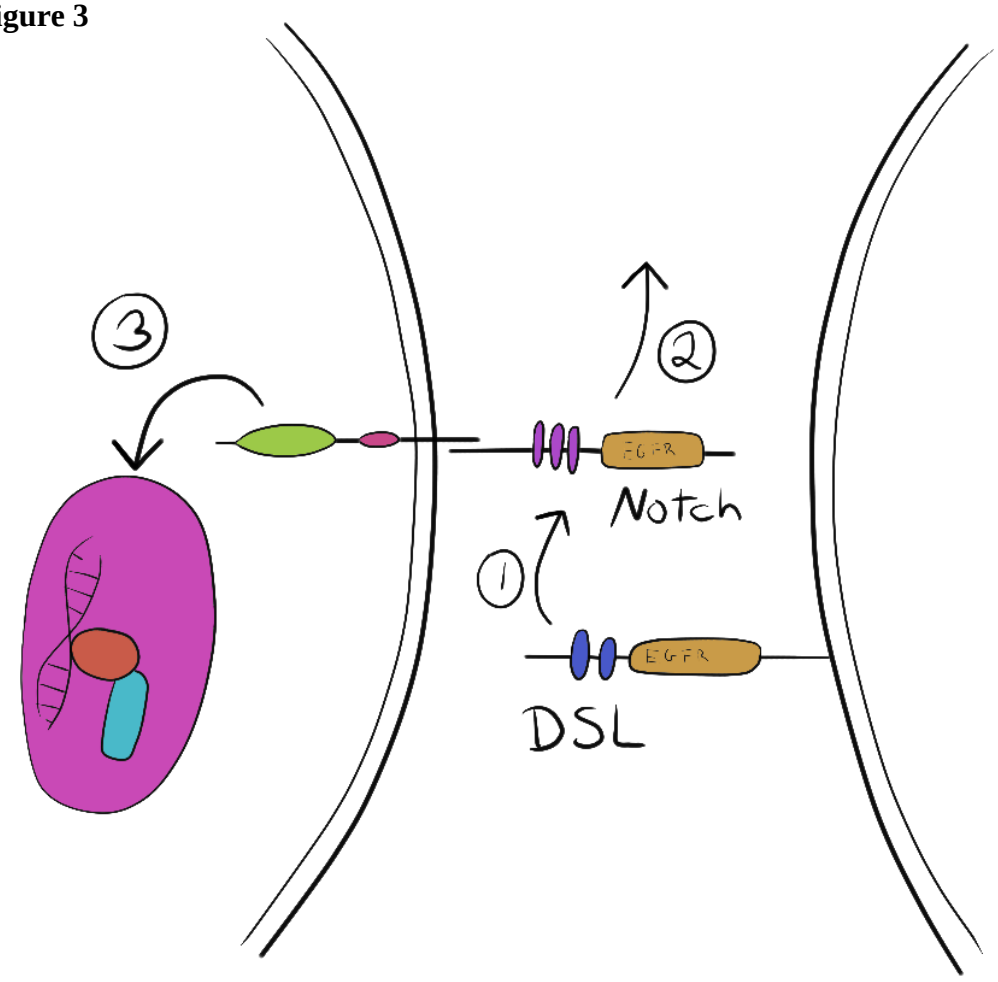


Figure 3. The Notch/DSL signaling pathway. 1) The DSL ligand contacts the Notch receptor on the receiving cell. 2) The extracellular part of the Notch protein is cleaved off, leading to a signaling cascade. 3) The cascade leads to the cleavage of the inner part of Notch, allowing it to then enter the nucleus and affect transcriptional changes.

Methods

Genomic database analysis

To examine pathway diversity, evolutionary trees (generated by TreeFam.org, see **Figure 4**) for various key genes in each pathway were collected and compared to each other. Trees were examined for each gene separately, and the number of branches for each species for each gene was counted. These counts were distilled into a table, and color-coded for ease of readability.

Figure 4

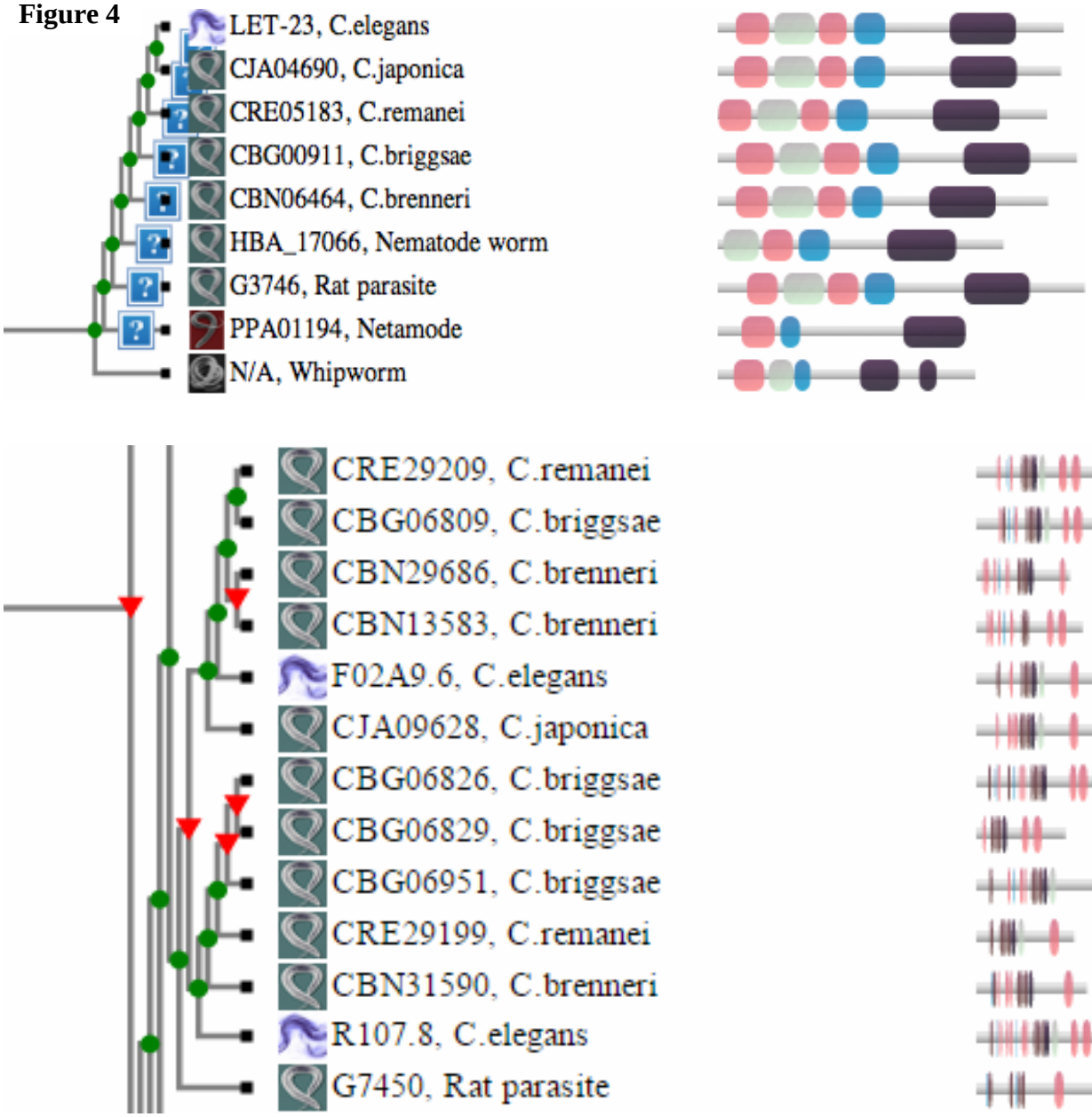


Figure 4. Example trees from TreeFam.org. The first tree is from the EGF/Ras pathway and shows *let-23*. The second tree is for the Notch signaling pathway and shows *lin-12* and g*lp-1*. These counts show one gene in the EGF/Ras path for this tree, but several duplications for the Notch/DSL tree, especially for *C. briggsae* and *C. brenneri*. The figures to the right represent protein domains.

***Isolation of genomic clones for* Cbr-lin-3 *and* Cel-lin-3**

To further study the EGF/Ras pathway, the signaling ligand gene *lin-3/EGF* was chosen for examination. Over expression of *lin-3/EGF* in *C. elegans* leads to the inappropriate division of epithelial cells. This results in the generation of several lumps on the ventral side of the worm, known as a Multivulva (Muv) phenotype (Sternberg, 2005).

To test if a similar result would occur in *C. briggsae* worms, an experiment was designed to overexpress the gene in wild-type (WT) *C. briggsae* worms. Wild-type AF16 worms were grown on NGM agar plates with E. coli OP50 bacteria at 20^{o} C, then harvested for DNA and RNA extraction. Genomic DNA was separated from RNA and extracted using Trizol and stored at -20^{o}C until use. The *lin-3* gene was isolated from the genomic DNA via Phusion PCR, using primers generated based on the *lin-3* sequence in WormBase. The PCR product was purified via column purification, and then digested with KpnI and SacI at 37^{o}C overnight. The blue-script vector pBSKII KS (+) was digested with the same enzymes, then dephosphorylated with Antarctic Phosphatase at 37^{o}C for one hour to help prevent the vector from closing on itself during ligation. Ligation was performed at room temperature overnight using T4 DNA Ligase. The ligations were transformed into DH5α *E. coli* cells, and grown overnight on LB+Carbenicillin (Carb) agar plates (50ug/mL) with IPTG and X-Gal. Colonies were checked for successful gene insertion via Taq PCR, then cultured in liquid LB+Carb overnight, for harvesting the following day via plasmid prep. The plasmids were verified via sequencing, then inserted into the gonads of AF16 worms via microinjection. Integration of the DNA into the genome resulted in a line of worms with multiple copies of the *lin-3* gene.

These worms were then observed using microscopy to see if they displayed a similar altered phenotype as *C. elegans* worms with multiple copies of the *lin-3* gene; control AF16

worms were also observed (n=30). Worms were scored as one of five "phenotypes" based on their appearance: 0, 1, ≥2, Dead, and DNC (Did not count). A score of zero indicated a wild-type worm; a score of greater-than-or-equal-to-two indicated a Muv phenotype. Scores of one and all dead worms are possible Muv phenotypes, but were considered less conclusive evidence and thus not counted into the true Muv category. Worms that were not counted, were simply not found on the plate on the day of observation; they most likely crawled off the plate. These DNC data were discarded.

Characterization of* lin-3 *isoform sequences

Further examination was done on the particular splicing of the *lin-3* transcripts in *C. elegans, C. briggsae,* and *C. remanei*. Hill and Sternberg (1992) examined the *C. elegans lin-3* gene in detail, documenting EGF repeats, the transmembrane domain, and a possible protease site, among other details, based partly on the findings of Flanagan, Chan, and Leder (1991) and Massagué (1990). Of particular note is the potential protease site, which would allow a soluble form of the *lin-3* ligand to be generated, such as what has been demonstrated with the Kit-ligand in mice (Flanagan, Chan, & Leder, 1991). A hypothesis, is that this protease site allows the worms to create both a membrane-bound and diffusible form of the LIN-3/EGF protein, which would allow the anchor cell to communicate with cells other than the P.6 cell, and would allow for production of a signal gradient.

To test the accuracy of the predicted protease site, wild-type N2 worms were grown on NGM agar plates with *E. coli* OP50 bacteria at 20^{o} C, then harvested for DNA and RNA extraction. RNA was extracted using Trizol and stored at -80^{o}C until use. Complementary deoxyribonucleic acid (cDNA) was created via reverse transcription using SuperScript III reverse transcriptase and random primers; the cDNA was stored at -20^{o}C until use. The *lin-3*

gene was isolated from the cDNA via Phusion PCR, using primers generated based on the *lin-3* sequence in WormBase. The PCR product was purified via column purification, and then digested with NheI and SacI at 37ºC overnight. The pPD96.48 (L2531) fire vector was digested with the same enzymes, then dephosphorylated with Antarctic Phosphatase at 37ºC for one hour. Ligation was performed at 15ºC overnight using T4 DNA Ligase. The ligations were transformed into DH5α *E. coli* cells, and grown overnight on LB+Carb agar plates. Colonies were checked for successful gene insertion via Taq PCR, then cultured in liquid LB+Carb overnight, for harvesting the following day via plasmid prep. The plasmids were verified via sequencing. Clones both containing and excluding the protease site were successfully generated.

The final part of this examination was to generate various *lin-3* cDNAs for members of the Caenorhabditis genus to examine possible variations in splicing across the genus. For the purposes of this experiment, *C. elegans*, *C. briggsae, C. brenneri, C. remanei,* and *C. japonica* were planned for examination, however this report only includes results from *C. elegans, C. briggsae,* and *C. remanei*. To do this, cDNA for each species was generated from RNA harvested from wild-type strains via reverse transcription, and various cDNA splice variants were generated via Phusion PCR and inserted into the pBSKII KS (+) blue-script vector, transformed into *E. coli,* and harvested for sequencing, all using the methods described previously. At the time of writing, six clones of *C. elegans,* three clones of *C. briggsae,* and two clones of *C. remanei* cDNA had been generated and sequenced.

Results

More gene duplication exists in the Notch/DSL signaling pathway than in the EGF/Ras pathway

Results of the examination of the evolutionary tree data revealed more gene duplication in the Notch/DSL pathway among members of the Caenorhabditis genus, while relatively little gene duplication exists in the EGF/Ras pathway. While the majority of genes in the EGF/Ras pathway has only one or two copies, some genes in the Notch/DSL path in some species had as many as 8 or 9 copies. The *dsl-1* ligand had the most duplication events among the genes examined, and *C. brenneri* has the most gene duplication among the members examined (see **Table 1**).

Over-expression of Cbr-lin-3 *in* C. briggsae *animals results in increased division of vulval cells and a Muv phenotype*

Current data show overexpression of *C. briggsae lin-3* causes an altered phenotype in *C. briggsae* worms: a Muv phenotype akin to what has been previously observed in *C. elegans*. Both species exhibit multiple vulval protrusions on their side, a trait defining the multivulval (Muv) phenotype. This indicates that both *C. elegans* and *C. briggsae* react similarly to an overexpression of the *lin-3* gene.

Figure 5

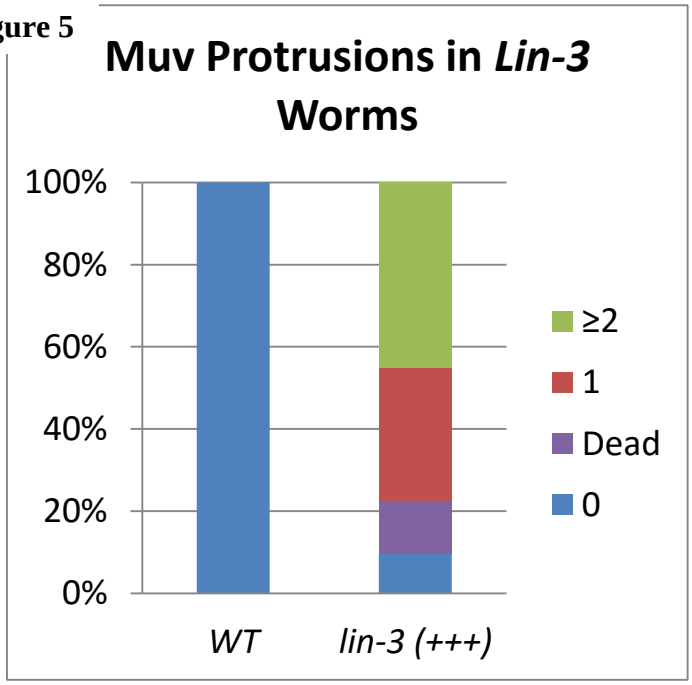


Figure 5. Muv data from experimental and control worms. Zero protrusions is characteristic of a WT phenotype; greater-than-or-equal-to-two protrusions is considered a true Muv phenotype. Other phenotypes (1, Dead) are considered to possibly be Muv as well, but are inconclusive.

Table 1

EGF/Ras		*C. elegans*	*C. briggsae*	*C. brenneri*	*C. remanei*	*C. japonica*
Ligand	***lin-3***	1	1	1	2	1
Receptor	***let-23***	1	1	1	1	1
Signal Transduction	***sem-5***	1	2	1	3	1
	sos-1 (let-341)	1	0	1	1	0
	let-60					
	lin-45	1	1	2	2	1
	mek-2	1	1	1	1	1
	gap-1	1	1	1	1	1
	gap-2	1	1	0	1	1
	gap-3					
Transcriptional Regulator	***mpk-1***	1	1	1	2	1
	lin-31					
	lin-1					
	eor-1					
	eor-2	1	1	2	2	1
	sur-2	1	1	1	1	0
	lin-25	1	1	2	2	1
Target Genes	***lin-39***					
	egl-17	1	1	0	1	0

Notch/DSL		*C. elegans*	*C. briggsae*	*C. brenneri*	*C. remanei*	*C. japonica*
Ligand	***lag-2***	1	1	1	1	1
	apx-1					
	dsl-1	8	5	7	9	3
Receptor	***lin-12/ glp-1***	2	4	3	2	1
Signal Transduction	***arg-1***					
	aph-1	1	1	2	1	1
	aph-2	1	1	1	1	0
	epn-1	1	2	5	2	1
	sup-17	1	1	2	1	1
	adm-4					
	hop-1/ sel-12	2	2	2	2	2
	pen-2	1	1	0	1	1
Transcriptional Regulator	***lag-1***	1	1	7	2	1
	sel-8	1	1	5	1	1
Target Genes	***ark-1***	1	1	1	1	1
	lip-1					
	lst-1					
	lst-2	1	1	1	1	1
	lst-3					
	lst-4	1	1	1	1	1

Table 1. Data collected from TreeFam.org, distilled into table format. Darker colors indicate more gene duplication. The Notch/DSL path is more darkly colored overall, indicating more duplication compared to the EGF/Ras path. Gened with no numbers listed (purple cells) had no tree data available, and were not included in the counts. Cells with a zero (0) had no copies listed in TreeFam for that gene for that species, which may be due to lack of data, or sequence divergence.

Gene structure and alternative splicing of Caenorhabditis **lin-3** *genes*

Examination of various *lin-3* cDNAs in *C. elegans*, *C. briggsae*, and *C. remanei* revealed several new splice variants of the gene. In *C. briggsae*, RNA-seq sequence data suggested the existence of two small exons upstream of the start of the gene predicted by the data available on WormBase. The results of bulk PCR sequencing – and later individual clone sequencing – revealed several *C. briggsae* cDNA reads that did in fact include one or more of these predicted exons, thus confirming their existence. Other sequences had a truncated version of exon 4, whose sequence is similar to the end of exon 5 and start of exon 6 in *Cel-lin-3*. This region is the site predicted by Hill and Sternberg in 1992 to be a potential protease site. Various *C. elegans* cDNAs, some not previously listed on WormBase were identified, including reads that encode proteins that both included and excluded the possible protease site.

Various cDNAs for *C. brenneri*, *C. remanei* were generated using PCR. Reads for *C. remanei* were successfully cloned into pBSKII KS (+) blue-script, transformed into *E. coli*, harvested, and sequenced. The reads obtained all had a portion of exon 4 missing, a region corresponding to the possible protease site in *C. elegans* and *C. briggsae*. All reads for each of the three species are documented in **Figures 6**, **7**, and **8**. Full CLUSTAL alignments for each species can be found in Appendices I, II, and III.

Clones of *Cel-lin-3* cDNA both encoding and excluding the potential protease site, under control of the *myo-2* promoter, were successfully generated, and injection mixes were made. These clones would express *lin-3* in the pharynx of the worms, away from the vulval cells. If the protease site is present, the protein product should be cleaved, and be able to diffuse to the vulva and cause a Muv phenotype; without the protease site, the LIN-3/EGF would not be able to

diffuse and act in the vulval cells. At time of writing, these mixes had not yet been injected into any worms, and thus no further data exists for this part of the experiment.

Figure 6

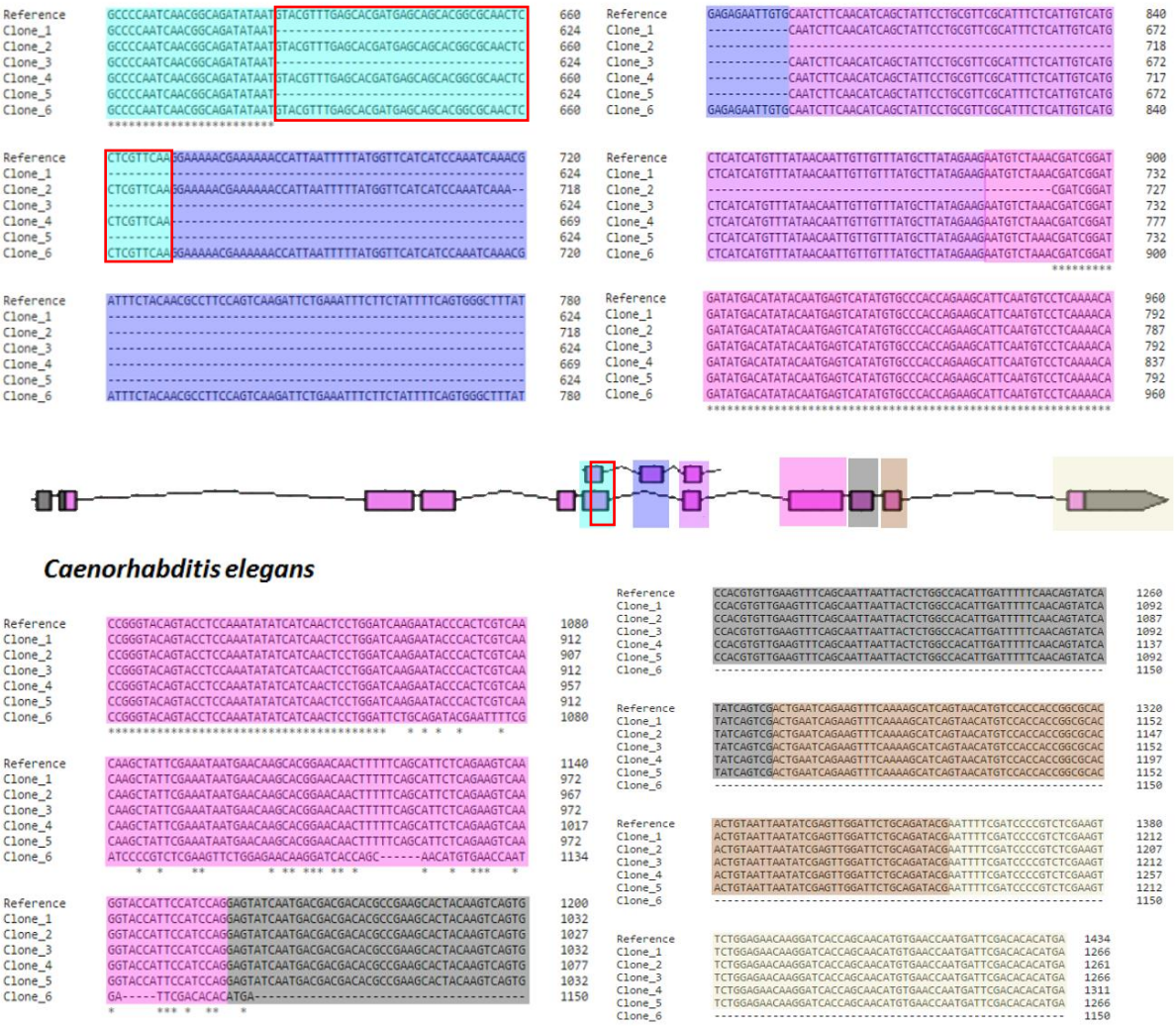


Figure 6. CLUSTAL alignment showing l*in-3* splice variants in *Caenorhabditis elegans*. Colors correspond to exons: Light Blue for Exon 5, Dark Blue for Exon 6, Purple for Exon 7, Pink for Exon 8, Dark Grey for Exon 9, Brown for Exon 10, and Light Tan for Exon 11. The red box represents the predicted protease site predicted by Hill and Sternberg (1992). Clone 6 is missing the end of Exon 8, and all of Exons 9, 10, and 11, thus it truncates early. The full CLUSTAL alignment can be found in Appendix I.

Figure 7

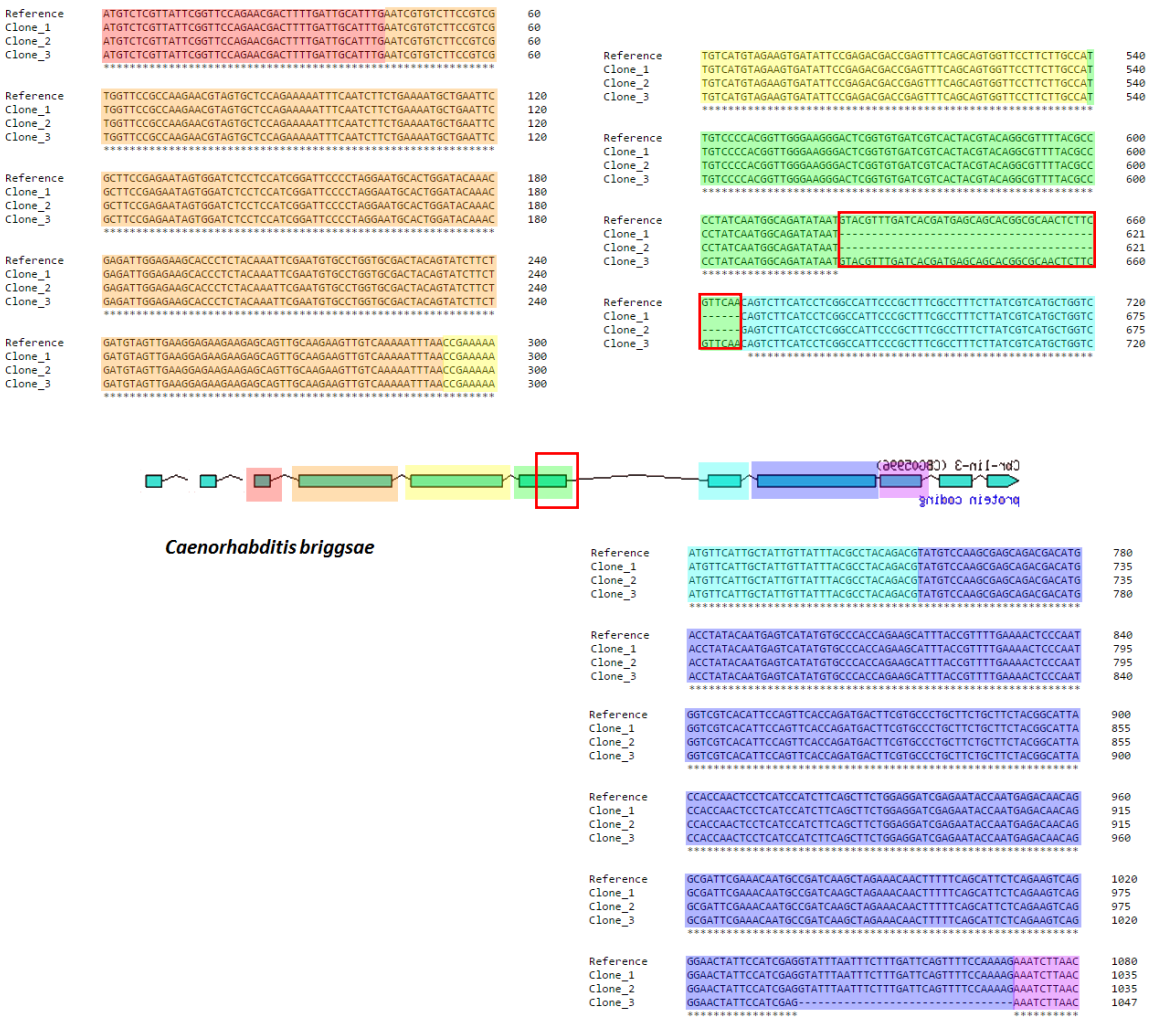


Figure 7. CLUSTAL alignment showing l*in-3* splice variants in *Caenorhabditis briggsae*. Colors correspond to exons: Red for Exon 1, Orange for Exon 2, Yellow for Exon 3, Green for Exon 4, Light Blue for Exon 5, Dark Blue for Exon 6, and Purple for Exon 7. The red box represents the predicted protease site. Two additional exons were discovered before Exon 1; these are not included in the CLUSTAL alignment, but are shown in the gene diagram above. The full CLUSTAL alignment can be found in Appendix II.

Figure 8

Figure 8. CLUSTAL alignment showing l*in-3* splice variants in *Caenorhabditis remanei*. Colors correspond to exons: Yellow for Exon 3, Green for Exon 4, Light Blue for Exon 5, and Dark Blue for Exon 6. The red box represents the predicted protease site. Two clones were analyzed, but their sequences were identical, thus they are represented by only one read: Clone_1. The full CLUSTAL alignment can be found in Appendix III.

Discussion

Evolutionary flexibility in EGF/Ras and Notch/DSL

The results of the examination of the evolutionary tree data revealed more gene duplication in the Notch/DSL pathway among members of the Caenorhabditis genus, while relatively little gene duplication exists in the EGF/Ras pathway. This suggests that the Notch/DSL pathway may be more evolutionarily flexible than the EGF/Ras pathway. Since gene duplication events likely occur in both pathways, but duplication seems to persist more or occur more frequently in components of the Notch/DSL pathway, this suggests that Notch/DSL is more tolerant of duplication, and thus alterations to the pathway. Since very little duplication is seen in EGF/Ras, it is likely that genes are not duplicated as frequently, or gene duplications are selected against evolutionarily.

One possible reason for this, is because the EGF/Ras signal sent by the anchor cell kicks off the process of division in the P6.p cell, and thus indirectly drives divisions for the P5.p and P7.p cells, as well as their neighbors. Since the EGF/Ras path is so crucial to the start of this process, it is possible that any changes to the path could result in severe detriment to the developing worm. Because Notch/DSL is only a lateral signal encouraging P5.p and P7p to divide (and since there is evidence EGF/Ras may play a part in inducing these cells as well), it has a less-strict role to play, and thus changes in the path do not necessarily result in catastrophic failure of vulva development.

Lin-3 ***splice variants that may affect processing are present in other Caenorhabditis nematodes***

The results obtained in the *lin-3* splicing examinations suggest that there are a variety of thus far undocumented splice variants of *lin-3* cDNA in *C. briggsae*, *C. elegans, and C. remanei*

worms, and possibly in other members of the Caenorhabditis genus as well. This suggests that there is greater variation in the protein products produced than what has been previously documented. This variation could help account for the lack of genome duplication events in the EGF/Ras pathway; diversity instead exists in splice variants instead of duplicated genes. As the worm develops, mRNAs are generated and alternatively spliced to provide what protein product the developing vulva needs at any particular time.

Future Work

Future plans include expanding the examination of *lin-3* cDNAs across other members of the Caenorhabditis genus. *C. brenneri* and *C. japonica* are current candidates, however there are many other members of Caenorhabditis that could be examined as well. Overexpression of *lin-3* in other Caenorhabditis members could also be an interesting test, to see if the Muv-response occurs in many members of the genus, or perhaps only after a certain point in the evolutionary branching. Once this information is obtained, an examination of changes in the *lin-3* gene across the evolution of the Caenorhabditis genus could provide information about how the EGF/Ras pathway has evolved over time.

There are also plans to inject wild-type *C. elegans* and *C. briggsae* worms with versions of *lin-3* cDNA under control of the *myo-2* promoter which expresses in the worms' pharynx. If versions with the predicted protease site cause a Muv phenotype, despite being expressed far away from the vulva, it would provide evidence that the protease site is accurate, and allows a diffusible form of LIN-3/EGF to be produced and travel to the vulva. This data would help support or refute the postulations of Hill and Sternberg, and would provide more information about how the EGF/Ras pathway operates to help generate a functioning vulva in nematode worms.

References

Greenwald I. and Kovall R. (January 17, 2013). Notch signaling: genetics and structure. *WormBook, ed. The* C. elegans *Research Community, WormBook,* doi/10.1895/wormbook.1.10.2, http://www.wormbook.org.

Flanagan, J. G., Chan, D. C., & Leder P. (March 8 1991). Transmembrane form of the *kit* ligand growth factor is determined by alternative splicing and is missing in the Sl^d mutant. Cell **64**, 1025 – 1035.

Hill, R. J. and Sternberg, P. W. (August 06 1992) The gene *lin-3* encodes an inductive signal for vulval development in *C. elegans*. Nature **358**, 470 – 476.

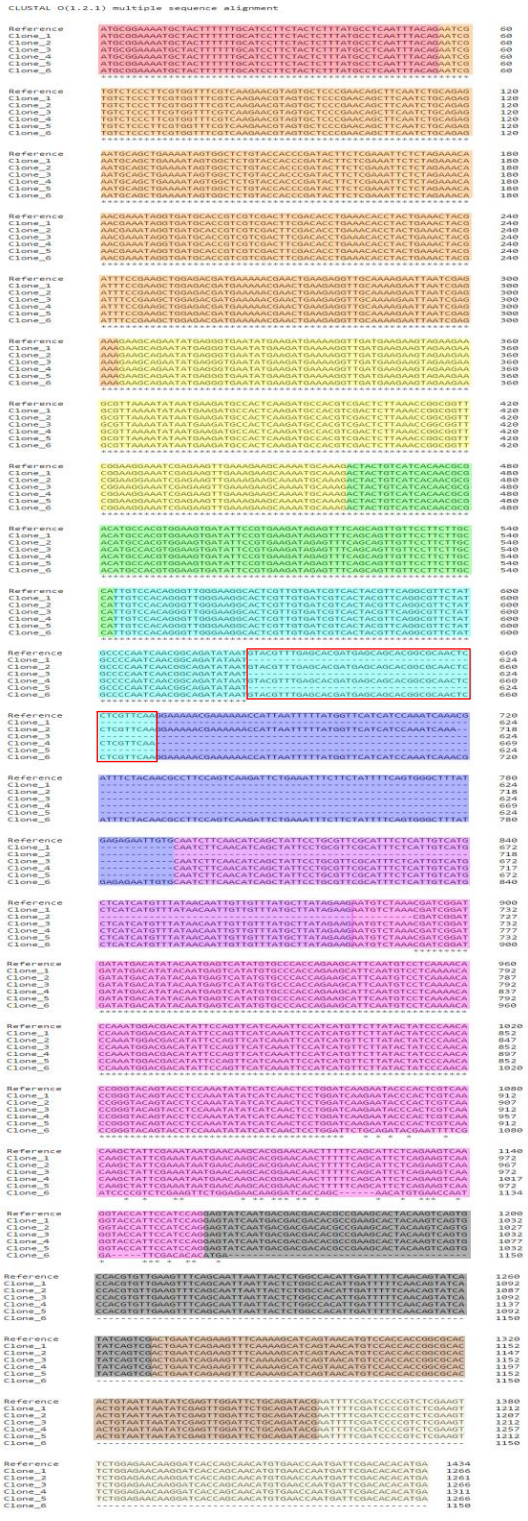
Massagué, J. (1990). Transforming growth factor-α: A model for membrane-anchored growth factors. J. Biol. Chem. **265**, 21393 – 21396.

Sundaram, M. V. (2005). The love-hate relationship between Ras and Notch. Genes Dev. **19**, 1825-1839.

Sundaram, M.V. (July 1, 2013). Canonical RTK-Ras-ERK signaling and related alternative pathways. *WormBook, ed. The* C. elegans *Research Community, WormBook,* doi/10.1895/wormbook.1.80.2, http://www.wormbook.org.

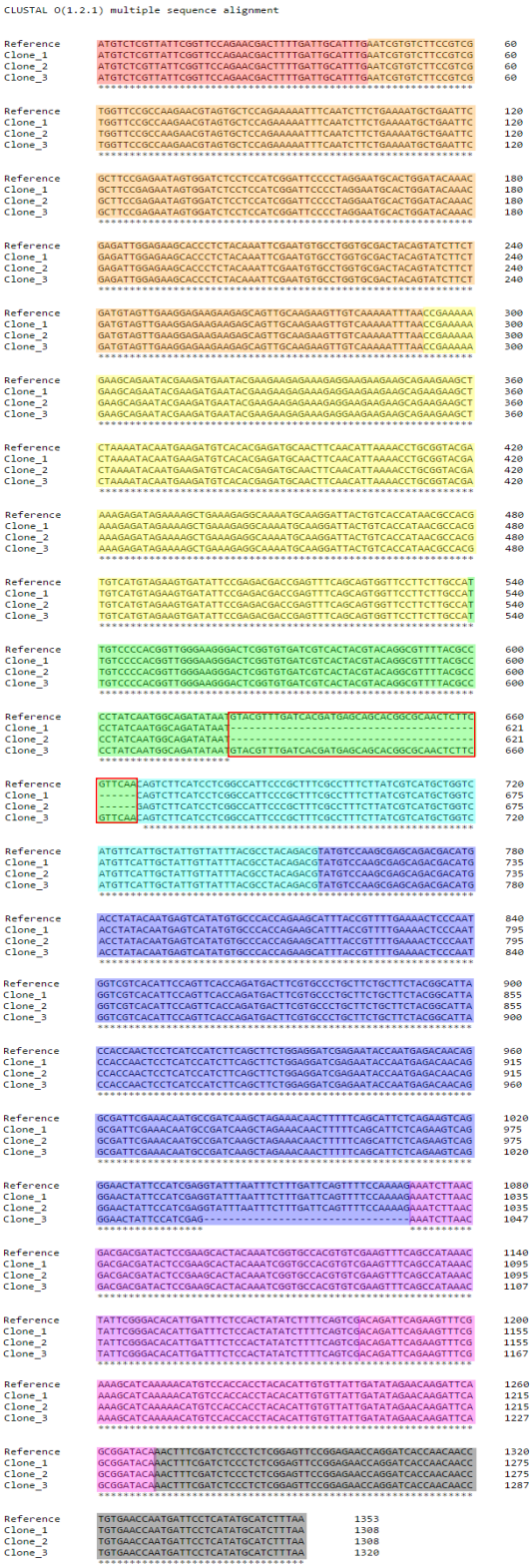
Sternberg, P.W. (June, 25 2005). Vulval development. *WormBook, ed. The* C. elegans *Research Community, WormBook,* doi/10.1895/wormbook.1.6.1, http://www.wormbook.org.

Appendix I. *Caenorhabditis elegans lin-3* CLUSTAL alignment*



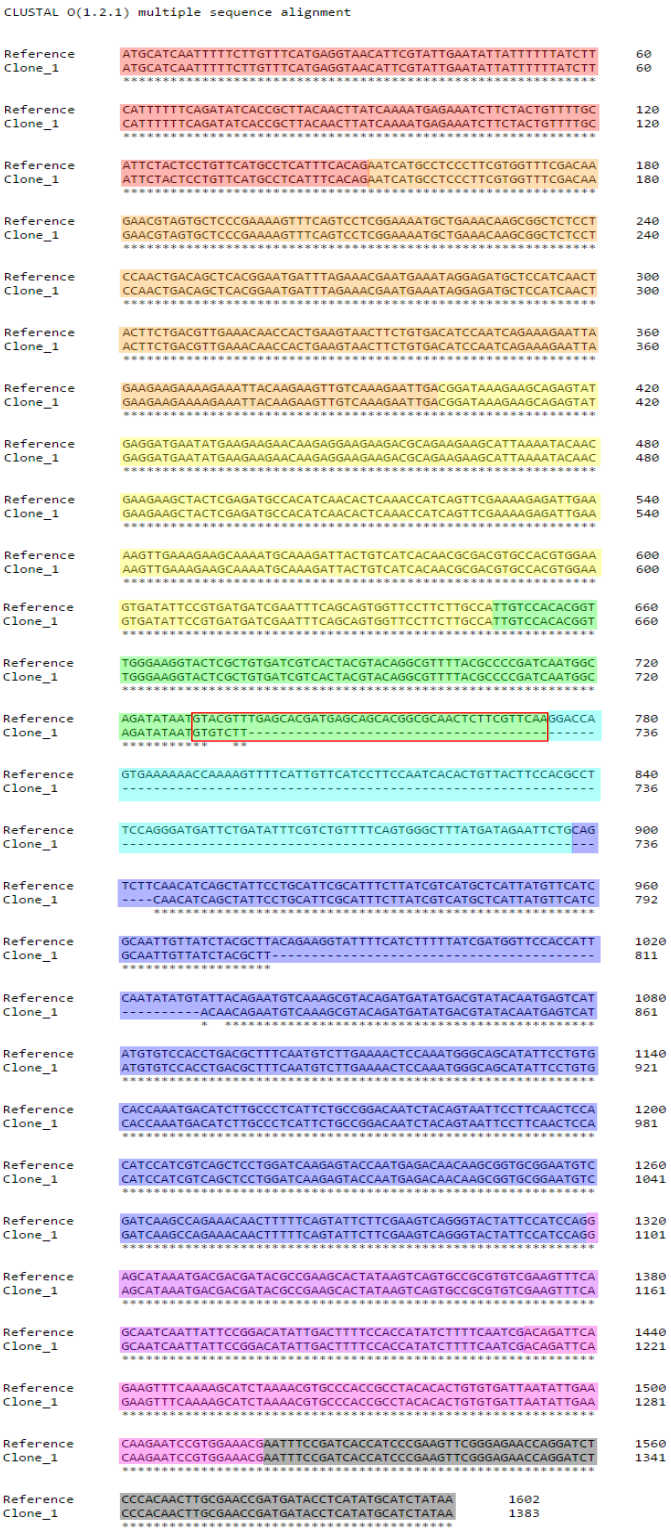
* Colors represent exons 1 through 11 in order, from red to light tan. The red box highlights the potential protease site.

Appendix II. *Caenorhabditis briggsae lin-3* CLUSTAL alignment**



** Colors represent exons 1 through 9 in order, from red to dark grey. The red box highlights the potential protease site.

Appendix III. *Caenorhabditis remanei lin-3* CLUSTAL alignment***



*** Colors represent exons 1 through 9 in order, from red to dark grey. The red box highlights the potential protease site.